



Modeling Postcolonial Language Varieties: Challenges and Lessons Learned from Mozambican Portuguese

Simone Ashby¹, Silvia Barbosa¹, Catarina Silva¹, Paulino Fumo², José Pedro Ferreira¹

¹Instituto de Linguística Teórica e Computacional (ILTEC), Lisbon, Portugal

² Faculty of Letters, Eduardo Mondlane University, Maputo, Mozambique

{simone, silvia, catarina, jpferreira}@iltec.pt, paulinofumo@yahoo.com.br

Abstract

While a growing number of studies deal with the modeling of accents by non-native speakers – with a strong bias towards English – relatively little is understood about postcolonial language varieties and the effects of substratal transfer (i.e. interference from local tongues) on the resulting sound systems. As the next information technology frontier [1], Africa presents unique challenges for language technologists tasked with adapting existing speech and language technologies (SLTs) to accommodate for this often systematic form of interference. As part of an effort to model and generate rule-based pronunciation lexicons for resource-scarce varieties, we present our work on the variety of Portuguese spoken in the Mozambican capital of Maputo. We highlight some of the segmental processes that result from contact with local Bantu languages, i.e. denasalization, aspiration, and rhotic variation. We then discuss these in the context of adapting Mozambican Portuguese (MP) to existing SLTs. Problems of inter- and intra-speaker variability and the lack of a well defined standard are also considered, along with implications for adopting a suitable model of MP for speech synthesis.

Index Terms: language modeling, speech varieties, resource-scarce, postcolonial, contact languages, pronunciation lexicon, substratal interference, bootstrapping, speech synthesis, Mozambique, Portuguese, southern Bantu, SLT4D, ICT4D

1. Introduction

MP is a postcolonial language that is in contact with a multiplicity of indigenous southern Bantu languages and dialects. The speech of L1 and L2 informants alike typically exhibits significant degrees of substratal transfer, indicating strong faithfulness constraints with regard to phonotactics, nasalization phenomena, voiced and aspirated stops, syllable shape, and rhythm [2,3].

In this paper, we describe the challenges encountered and lessons learned from modeling the variety of MP spoken in Maputo, Mozambique, hereafter referred to as P_{Map} . Our analysis of P_{Map} is particularly focused on the effects of local Bantu substrates on informant realizations of read speech, and implications for adapting existing Portuguese SLTs to P_{Map} and other Luso-African varieties. We also probe the issue of a lack of a well defined MP standard, and the implications this has for selecting an appropriate model for TTS, given the considerable amount of variability observed in our corpus.

We conclude by broadening the discussion to other postcolonial varieties, offering a linguistic perspective on how best to approach the bootstrapping of SLTs – and TTS systems in particular – in resource-scarce environments.

2. Background

2.1. The SLT4D initiative

Regionalized varieties of English, French, and Portuguese are widespread in sub-Saharan Africa and offer a means of bridging the digital divide that exists between developing nations of Africa and the rest of the world through the adaptation of existing SLTs. This work fits within the Speech and Language Technologies for Development (SLT4D) initiative, involving, among other things, the development of speech technologies for resource-scarce languages and varieties [4].

2.2. Why bring SLTs to developing regions?

SLT4D is part of the larger Information Communication Technologies for Development (ICT4D) initiative, which calls for technological solutions to problems in underserved regions [5]. This aim is well reflected in the deployment of mobile-based SLTs in developing regions, where illiteracy rates are high, and where populations have little access to computers and the Internet [4]. Such interfaces have already proven useful in Pakistan through the provision of a telephony-based information service for low-literate health workers [6], and by providing low-skilled, low-literate people in Lahore with a spoken dialogue system for browsing job ads with their mobile phones [7].

2.3. Lack of linguistic resources

A significant hurdle in the development of robust SLT4D applications concerns the lack of key linguistic resources upon which SLTs are based, such as electronic dictionaries, phonetically aligned corpora, pronunciation lexicons, and detailed published studies. Pronunciation dictionaries are an important component in the development of localized TTS and ASR systems, among other SLT implementations. However, these are typically very expensive to produce.

For postcolonial language varieties, large-vocabulary pronunciation lexicons designed for other user populations present a way forward. Yet, without sufficient efforts to adapt these lexicons, ASR performance will be degraded, e.g. [8]. Also, target users may likely find it difficult to warm to a spoken dialogue system that lacks cultural adaptation in terms of the language variant used, e.g. [9].

2.4. LUPo and the Unisyn methodology

The work described in this paper was conducted for the LUPo (Portuguese Unisyn Lexicon) project, dedicated to: (1) modeling standard and non-standard varieties of spoken Portuguese; (2) delivering a free, open-source tool for the automatic generation of accent-specific pronunciation lexica via the existing *Portal da Língua Portuguesa* knowledge base

(<http://www.portaldalinguaportuguesa.org>); and (3) providing the R&D community with a free, online, searchable databank concerning the description and modeling of regional varieties of spoken Portuguese [10,11,12].

The Unisyn model upon which LUPo is based [13,14,15] offers a unique approach to modeling pronunciation variation through investment in an accent-independent lexicon and rule system for generating high quality accent-specific phonetic output. The Unisyn framework offers a robust method for generating large-scale pronunciation dictionaries for multiple varieties and handling out-of-vocabulary items. It is also well suited for representing varieties with limited linguistic resources, provided an initial investment is made involving two or more highly trained linguists.

Our methodology involves a careful modeling of the accent’s sound system. This information is interpreted as a set of hand-written rules, which are applied to an accent-independent lexicon (i.e. a list of words with their X-SAMPA-encoded metaphonemic representations) for generating accent-specific phonetic transcriptions. Phonetic conversions are governed by a geographically organized regional accent hierarchy that enables the sharing of rules across regional varieties. In this way, LUPo – like its English language precursor – succeeds in dramatically reducing the investment spent per regional variety, while yielding high-quality pronunciation output. Introducing new rules for P_{Map} , for example, facilitates the work of adding additional regional varieties from Mozambique and other Luso-African countries that share a common store of allophones. This is very meaningful for adapting rule systems to resource-scarce varieties and establishing tailor-made pronunciation dictionaries in a relatively low-cost manner.

3. Creating a model of P_{Map} ¹

3.1. Corpus

The analysis that follows is based on the read speech of six university educated male and female informants from Maputo. As indicated in Table 1, the informants range in age from 21-33, five of whom regard Portuguese as their first language, while Informant 047 reported learning Portuguese during the early years of her primary education. In addition, informants reported varying degrees of fluency in the Changana dialect of Tsonga, as well as Ronga, Zulu, English, and French.

ID	037	039	042	043	045	047
Age	26	33	21	27	27	26
Sex	M	M	F	F	M	F
Birthplace	Maputo	Maputo	Maputo	Maputo	Maputo	Maputo
MP is L1 or L2	L1	L1	L1	L1	L1	L2
Other languages spoken fluently	Tsonga, French, English	Tsonga, English, Zulu	Tsonga, Ronga	Tsonga, English	Tsonga, Ronga	Tsonga, French
Mother’s L1	Tsonga	Tsonga	Portuguese	Tsonga	Tsonga	Tsonga
Father’s L1	Tsonga	Tsonga	Portuguese	Tsonga	Chope	Tsonga

Table 1: *Informant details.*

¹ Note that in the transcriptions that follow, ‘.’ represents a syllable boundary and ‘*’ indicates elision.

For the read speech elicitation task,² informants were asked to read the individual phrases and sentences projected in front of them on PowerPoint slides. Recordings were performed with a Marantz digital voice recorder.

3.2. Methodology

Corpus-based accent models were developed at the segmental level by first establishing a list of contexts in which the metaphonemic representations described in Section 2.4 were likely to have a different surface realization for the variety being modeled. This was accomplished by listening to read-speech data for three of the L1 informants, creating a new list entry when alternate surface realizations were heard, and noting the observed realizations in the form of rules. For example, the plural suffix <-ões>, represented in LUPo’s accent independent lexicon as |-o~j~s|, is a context for the MP rule **do_oj_oi_resyllabify**, which converts the oral or nasal diphthong [oj] to the resyllabified monothong sequence [o.i].

The list of contexts was then expanded and modified using observations from data for the remaining informants, with a focus on those contexts which yielded variable output in the original three informant recordings analyzed. In cases where a metaphonemic context resulted in variable output, frequencies were calculated across informants, with an indication of the most stable variant per context.

By accent, we mean those features characteristic of a speaker’s pronunciation. While future plans involve the modeling of suprasegmental characteristics, our current efforts are focused on describing the broad segmental features of P_{Map} . In terms of style, this model reflects a slower and more carefully articulated variety than that found in conversational speech.

3.3. Transfer from the Bantu substrate

The effects of local Bantu substrates on informants’ spoken Portuguese varieties are often attested in the literature on Bantu linguistics, and have been modeled in the form of rules. E.g. the rule **do_aspirant_resyllabify** for the realization of substitutive aspiration and vowel apocope in unstressed, word-final contexts, such as be[b^{h*}] *bebo* ‘I drink’, and objeti[v^{h*}] *objetivo* ‘objective’, realized respectively in European Portuguese (EP) as be[bu] (or be[βu]) and objeti[vu].

In the sections that follow, we offer a qualitative analysis of these and other phenomena. Where necessary, examples are also provided in EP as a means of presenting a baseline for comparison in the identification of Bantu substratum interference in P_{Map} .

3.3.1. Denasalization

The nasalized vowel context in Portuguese frequently undergoes denasalization in P_{Map} in the context of a following continuant, and for some informants before a heterosyllabic voiceless consonant. We interpret this phenomenon as resulting from the following premises:

- Prenasalized voiced and voiceless stops are a common feature in the southern Bantu languages, e.g. [16].

² The elicitation of spontaneous speech data was also conducted in the form of an oral questionnaire for obtaining general speaker information and attitudinal data.

- P_{Map} speakers regularly produce the Portuguese heterosyllabic sequence: nasal vowel (V_n) + stop (C_s) + V. Note, however, that this does not apply in the case of clusters – e.g. eng[e*.dr]ou (EP: eng[ẽ.dr]ou) *engendrou* ‘engendered’ and c[o*.tr]ariedade (EP: c[õ.tr]ariedade) *contrariedade* ‘setback’ – in which the preceding nasal mora is typically elided by P_{Map} speakers.
- Relatively few tokens of the Portuguese sequence V_n + continuant (C_c) were observed, indicating that V_n + C_s sequences may perhaps be regarded by P_{Map} speakers as prenasalized stops, while V_n + C_c sequences are dispreferred and substituted with a corresponding V + C_c sequence (see also [2,3]).

The above phenomena are illustrated in (1), which shows the permissibility of V_n + C_s sequences (1a), (1b), (1c), (1d), and a relative dispreference for V_n + C_c sequences (1e), (1f).

(1) P_{Map}	EP	Informant
a. Orl[ã̃n]do ‘Orlando’	[ã̃]	*all*
b. [ã̃ŋ]gústia ‘anguish’	[ã̃]	*all*
c. so[õ̃m]breado ‘shaded’	[õ̃]	*all*
d. c[o*]plexa ‘complex’	[õ̃]	042, 043
e. Ev[a*]gelho ‘Gospel’	[ã̃]	*all but 047*
f. g[o*]zo ‘hinge’	[õ̃]	042, 043

Denasalization in P_{Map} is also a frequent occurrence in the aforementioned plural suffix <-ões> (EP: [-õj]), which is often realized as the monothongized [-o.ɐf], e.g. in words such as *delações* ‘vigilantism’ and *populações* ‘populations’. The same may be observed in the 3P-sg. verbal suffix and nominalizing diphthong <-em> (EP: [-ẽj]), which is typically realized as [-ej], e.g. in words such as *tem* ‘he has’ and *ferrugem* ‘rust’.

3.3.2. Aspiration

The use of sustained aspiration is extremely common in P_{Map} . Aspirated voiceless stops contrast with their unaspirated counterparts in a variety of southern Bantu consonantal inventories, e.g. [16,17]. The aspiration of consonants when followed by a high vowel is also attested in [17] for languages such as Makhuwa and Doko. In our own corpus, we observed similar forms of aspiration, e.g. in epilé[t^hi.k^hu] (EP: epilé[ti.ku]) *epiléptico* ‘epileptic’ and prínci[p^hi.u] (EP: prínci[pju]) *princípio* ‘principle’. Sustained aspiration also serves as a ‘filler’ for elided word-final unstressed vowels. In this context, it is possible that aspiration serves as a type of rhythmic counterbalance, in that it is always preceded by a heavy stressed vowel.

Observe the following examples:

(2) P_{Map}	EP	Informant
a. unicamen[t ^h *] ‘only’	[t̥i]	*all*
b. saio[t ^h *] ‘kilt’	[t̥i]	*all*
c. estupefac[t ^h *] ‘thunderstruck’	[tu]	*all but 037, 043*
d. Trinda[d ^h *] ‘Trinity’	[d̥i]	*all*
e. esperança[d ^h *] ‘hopeful’	[du]	039 ¹
f. be[b ^h *] <i>bebo</i> ‘I drink’	[bu]	042, 047
g. objeti[v ^h *] <i>objetivo</i> ‘goal’	[vu]	039, 047

¹ Informant 037 produced a homorganic epenthetic nasal [n] before /-du/, thus removing the context for substitutive aspiration. Somewhat similarly, Informant 042 produced a geminated [dd], while Informants 043 and 045 produced tokens with stopping and fronting [dð], thereby necessitating preservation of the final vowel (see also [2,3]).

The examples in (2) show word-final aspiration in the context of an elided word-final unstressed vowel. Consistent use of this type of aspiration can be seen in the examples with /t/ + an unstressed vowel, i.e. (2a), (2b), and (2c). We also see evidence of word-final aspiration occurring after /d/ – e.g. (2d) and (2e) – although see the previous footnote concerning (2d). Note also the use of aspiration after /b, v/ in (2f) and (2g).

Figure 1 shows the realization of *estupefacto* for Informant 042, with an elongated period of aspiration substituting for unstressed, word-final /u/. Note the appearance of vowel-like resonances, possibly due to front and back cavity coupling as the informant maintained constriction in the front part of the vocal tract [18].

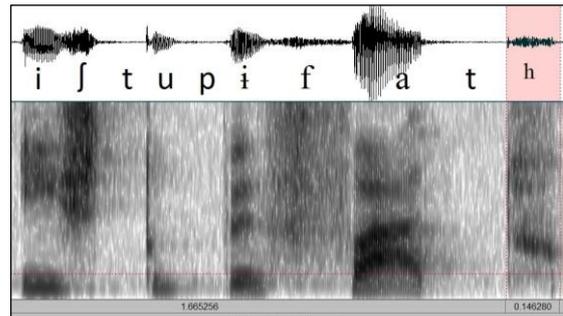


Figure 1: Waveform and spectrogram for *estupefacto* (042).

The sustained quality of substitutive aspiration in P_{Map} , and the consistency with which it may be observed both within and across speakers, shows the convergence of substratal rhythmic constraints with Bantu-like syllable shape preferences, whereby [ʰ] (or [h], as it should perhaps be labeled) assumes a vowel-like identity at the right edge of a word.

3.3.3. Rhotic Variation

Rhotic variation is one of the more interesting phenomena observed in P_{Map} and plays a role in adhering to the preferred CV syllable structure that is characteristic of the Bantu family. According to [19], the local Changana variety has /r/, which by other accounts is complemented by a phonologically distinct aspirated rhotic sound, spelled <rh> [20,21].

In P_{Map} , we note that informants use varying realizations of /r/, depending on the surrounding context. While some realizations correspond with EP, as well as Brazilian Portuguese (BP) – e.g. the consistent realization of [r] in non-nasal, intervocalic contexts, e.g. bebe[r]am *beberam* ‘they drank’ – others reveal strategies on the part of speakers for producing vocalic or near-vocalic morae in the context of a neighboring consonant.

Observe the following examples:

(3) P_{Map}	EP	Informant
a. transfo[r]mação ‘transformation’	[r]	037, 042, 043
b. Setemb[ə.r]o ‘September’	[r]	037, 043, 045
Setemb[r ^h]o		039, 042
c. câib[r]a ‘cramp’	[r]	037, 043, 045
câib[ə.r]a		042
d. escola[.r ^h] ‘academic’	[r]	*all*
e. gen[r]o ‘son-in-law’	[ɾ]	*all but 037, 039*

As demonstrated in Examples (3a), (3c), and (3e), informants produced an apico-alveolar trill in the context of a neighboring oral or nasal obstruent. What is particularly

noteworthy about these trills is the presence of a vowel-like transition bordering the neighboring obstruent.

Figure 2 shows the realization of /r/ in *cãibra* by Informant 037, in which vowel-like formants appear at the start of the trill (after the release of the stop closure for /b/), a type of phenomenon that has also been described for Greek clusters in [22]. The second example in (3c) – in which Informant 042 produces an intrusive schwa + trill – is a more exaggerated manifestation of this form of substratal interference.

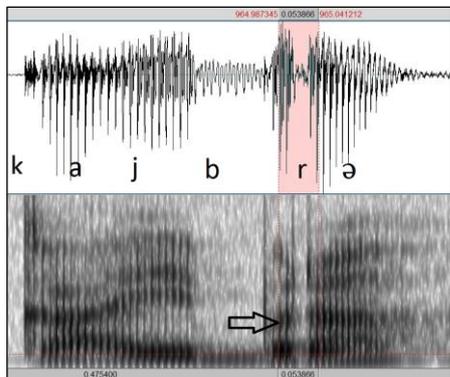


Figure 2: *Waveform and spectrogram of cãibra (037).*

Example (3b) shows an onset cluster for *Setembro* in which, by EP phonological accounts, the voiced stop /b/ is preceded by the nasal vowel /ẽ/. For MP speakers, this orthographic sequence is commonly interpreted as the prenasalized stop /^mb/ [3] – evidence of which may be seen in the realization of intrusive schwa + flap by informants 037, 043, and 045, versus the realizations [r] and [ə.r] for the nasal vowel contexts *cãibra* and *genro*. Again, this reflects an attempt by speakers to adhere to the preferred Bantu CV syllable shape, as well as remaining faithful to substrate rhythmic constraints governing syllable weight.

Example (3d) shows further evidence of the use of word-final aspiration in P_{Map} as a rhythmic counterweight (see Section 3.3.2), as well as a strategy for maintaining the optimal syllable shape. It also goes some way in validating the observance of an aspirated rhotic in Changana/Tsonga, which is documented in [20,21] and disputed in [19].

4. Variability and the lack of a standard

Variability must be accounted for in the treatment of any regional variety. However, given the degree of multilingualism described in Section 1, and the fact that the Portuguese language is still in the process of becoming nativized in Mozambique, phonetic variability should be considered a defining feature of P_{Map} , thus requiring adequate handling.

Some of the contexts in which P_{Map} is particularly prone to yielding variable output have been described in Section 3. One also frequently finds the non-systematic pronunciation of “silent” consonants /p/ and /k/ in words such as ado[**p**]çãõ *adopção* ‘adoption’, and da[**k**]tilografia *dactilografia* ‘typing’, which are no longer phonetically represented in BP and EP; along with the realization of /t/ as a stop [t] or a homorganic fricative [θ]; realization of the stressed nasal vowel /ẽ/ as a diphthong [ɛ̃] or a raised segment [ɨ]; and variable realizations of the open and close vowel pairs /ɛ, e/ and /ɔ, o/.

We applied informant frequencies in our decision making concerning the modeling of P_{Map} . Having collected detailed

background information from our informants, we are also making use of LUPo’s regional accent hierarchy (see Section 2.4) to attribute certain allophonic alternations to a specific speaker profile by introducing rules at the sub-TOWN level that define the characteristics of a particular city zone or neighborhood. In general, while our methods for modeling regional varieties through direct contact with speakers introduces a level of variability that may be considered unwieldy, we regard variability as essential in creating rules for generating high-quality pronunciation output. By cataloging all observed variants with their context in LUPo’s free, online, databank, we achieve the added goal of establishing a separate resource for use by other linguists and speech technologists.

The lack of a well defined standard for MP was perhaps best addressed, however, by LUPo’s native speaker consultants, who, when presented with pronunciation output reflecting our corpus-based rules, gave meaningful feedback on realizations that triggered some notion of incorrectness, or negative sentiment. It is of further note that this feedback, provided by university students in linguistics, showed little dialectal accommodation with the dominant EP and BP varieties.

5. Implications for bootstrapping

In terms of bootstrapping ASR applications to P_{Map} and other Luso-African varieties with Bantu interference, this work highlights the need for investment in local speech resources that not only inform the phone set and serve in the learning of pronunciation rules, but also provide linguistically motivated confidence measures for dealing with the variability that comes from substratal interference. LUPo’s online databank [10] is designed to facilitate these objectives by linking recorded data with both pronunciation rules and informant frequency and background information, so that this information might serve in the weighting of finite state transducers.

For TTS systems, the problem lies in choosing a locally accepted spoken variety and deciding which features to represent in the pronunciation lexicon. Here, the careful use of native speaker feedback is essential in gleaned local preferences, where no other standard exists. Rule-based hierarchical approaches such as the Unisyn methodology further enable the low-cost delivery of *multiple* local pronunciation lexicons suiting target audiences, e.g. Bantu-language defined communities. Such approaches also reduce the time and effort required for modeling additional Bantu influenced varieties, e.g. the densely populated region of Nampula (MZ), or the Portuguese of Angola.

6. Concluding remarks

Substratal transfer presents a significant source of variability that should be accounted for in the development of robust models for recognition and synthesis. With respect to resource-scarce, postcolonial varieties, bootstrapping methods require an investment in local speech resources that can assist in properly weighting this data, as well as informing the selection of features for use in locally acceptable TTS systems.

7. Acknowledgements

The authors gratefully acknowledge the support of the Fundação para a Ciência e a Tecnologia, and the cooperation of Dr. Susan Fitt, whose development of the Unisyn Lexicon is the inspiration for this work. We are also grateful to Inês Machungo and colleagues at the Eduardo Mondlane University.

8. References

- [1] The Economist, "The Next Frontier", 16th February 2013. Online: <http://www.economist.com/news/business/21571889-technology-companies-have-their-eye-africa-ibm-leading-way-next-frontier>.
- [2] Ashby, S., Barbosa, S., "Bantu substratum interference in Mozambican Portuguese Speech Varieties", *Africana Linguistica* 17:3-31, 2011.
- [3] Ashby, S., Barbosa, S., Silva, C., Ferreira, J. P., Fumo, P., "Syllable Dynamics in Mozambican Portuguese", abstract under review. Submitted to *Colonial and Postcolonial Linguistics* on 30 April 2013.
- [4] Sherwani, J., Rosenfeld, R. "The Case for Speech Technology for Developing Regions", Proc. Human-Computer Interaction for Community and International Development workshop, 2008. Online: www.cs.cmu.edu/~jsherwan/pubs/case-slt4d.pdf.
- [5] Brewer, E., Demmer, M., Du, B., Fall, K., Ho, M., Kam, M., Nedeveschi, S., Pal, J., Patra, R., Surana, S., "The Case for Technology for Developing Regions", *IEEE Computer* 38(6):25-38, 2005.
- [6] Sherwani, J., Ali, N., Mirza, S. Fatma, A., Memon, Y., Karim, M., Tongia, R., Rosenfeld, R., "HealthLine: Speech-based Access to Health Information by Low-literate Users", Proc. IEEE/ACM International Conference on Information and Communication Technologies and Development, 2007. Online: <http://www.cs.cmu.edu/~jsherwan/pubs/ictd07.pdf>.
- [7] Raza, A. A., Haq, F. U., Tariq, Z., Pervaiz, M., Razaq, S., Saif, U., Rosenfeld, R., "Job Opportunities through Entertainment: Virally Spread Speech-Based Services for Low-Literate Users", Proc. ACM SIGCHI Conference on Human Factors in Computing Systems, 2013. Online: <http://www.cs.cmu.edu/~Polly/pubs/CHI13.pdf>.
- [8] Rouas, J.-L., Trancosco, I., Viana, C., and Abreu, M., "Language and Variety Verification on Broadcast News for Portuguese", in *Speech Communication*, 50(11-12):965-979, 2008.
- [9] Pucher, M., Schuchmann, G., Fröhlich, P., "Regionalized Text-to-Speech Systems: Persona Design and Application Scenarios", in Esposito, A., Hussain, A., Marinaro, M., Martone, R. [Eds.], *Multimodal Signals: Cognitive and Algorithmic Issues*, LNAI 5398:216:222, Berlin-Heidelberg: Springer Verlag, 2009.
- [10] Ashby, S., Barbosa, S., Brandão, S., Ferreira, J. P., Janssen, M., Silva, C., Vilaro, M. E., "A Rule Based Pronunciation Generator and Regional Accent Databank for Portuguese", Proc. *Interspeech*, paper #160, 2012.
- [11] Ashby, S., Ferreira, J. P., Barbosa, S., Silva, C., "Rule Based Pronunciation Modeling for Resource-Scarce Varieties", in Proc. *Iberspeech*, 2012. Online: <http://iberspeech2012.ii.uam.es>.
- [12] Ashby, S., Ferreira, J. P., and Barbosa, S., "Adapting the Unisyn Lexicon to Portuguese: Preliminary Issues in the Development of LUPo", in Proc. *Iberian SLTech*, 43-46, 2009.
- [13] Fitt, S., "Documentation and User Guide to UNISYN Lexicon and Post-lexical Rules", Online Technical Report, Centre for Speech Technology Research, 2000. Online: <http://www.cstr.ed.ac.uk>.
- [14] Fitt, S., Isard, S., "Representing the environments for Phonological Processes in an Accent-independent Lexicon for Synthesis of English", Proc. *ICSLP*, 847-850, 1998.
- [15] Fitt, S., Isard, S., "Synthesis of Regional English Using a Keyword Lexicon", Proc. *Eurospeech*, 823-826, 1999.
- [16] Maddieson, I., "The Sounds of the Bantu Languages", in D. Nurse, Philippson, G. [Eds.], *The Bantu Languages*, New York: Routledge, 15-41, 2003.
- [17] Hyman, L. M., "Segmental Phonology", in D. Nurse, Philippson, G. [Eds.], *The Bantu Languages*, New York: Routledge, 42-58, 2003.
- [18] Stevens, K. N., "On the Quantal Nature of Speech", *Journal of Phonetics* 17:3-45, 1989.
- [19] Janson, T., "Consonants in Changana/Tsonga", *South African Journal of African Languages* 21:1, 16-32, 2001.
- [20] Cuénod, R., *Tsonga-English Dictionary*, Johannesburg: Swiss Mission in South Africa, 1967.
- [21] Siteo, B., *Dicionário Changana-Português*, Maputo: Instituto Nacional do Desenvolvimento da Educação (INDE), 1996.
- [22] Baltazani, M., "Acoustic Characterization of the Greek Rhotic in Clusters", Proc. 18th International Symposium on Theoretical and Applied Linguistics, 87-95, 2009.